Part 12: Multi-Fidelity Gaussian Processes

February 15, 2021

This should hopefully be an interesting post. We will be talking about using Gaussian Processes (GPs) for modeling with multiple data sources, or fidelities. For example, what is the true concentration of a pollutant if we have multiple measurements from different devices, given that said devices may have different biases, variances, locations, and costs of operation?

1 The Multi-Fidelity GP Model

Let us utilize fidelity levels l. We wish to model a true underlying process $f_0(x) \sim GP(0, K_0)$. We will utilize multiple fidelities $f_l(x) = g(x) + \delta_l(x)$ with a GP model of the deviance from the true process $\delta_l(x) \sim GP(0, K_l)$. The underlying process $l = 0$ can have a squared exponential covariance kernel:

$$K_0(x, x') = \sigma_f^2 exp(-0.5\Sigma_i^p \frac{(x_i - x_i')^2}{\lambda_i})$$

Which means the underlying process has $p+2$ hyperparameters. Due to $f_l(x) = g(x) + \delta_l(x)$ being the model of the fidelities $l > 0$, each process l will have it's own squared exponential kernel plus the underlying kernel:

$$K_l(x, x') = K_0(x, x') + \sigma_f^2 exp(-0.5\Sigma_i^p \frac{(x_i - x_i')^2}{\lambda_i})$$

We are also going to assume fidelity independence, which means $K_l(x, x')$ will be calculated as above only when $l = l'$ for x and x'. If $l \neq l'$ or $l = 0$ (as stated above) then $K_l(x, x') = K_0(x, x')$. So (as in the code attached) when calculating

the kernel matrix just use a bunch of *if* statements to construct the kernel based on the identity of the data x and x'.

To calculate mean and standard deviation:

$$\mu(x^*) = K(x^*, X)(K(X, X) + sI)^{-1}Y \tag{1}$$

$$\sigma^2(x^*) = K(x^*, x^*) - K(x^*, X)(K(X, X) + sI)^{-1}K(X, x^*) \tag{2}$$

Where s is a vector of $\sigma^2(X)$ hyperparameters.

2 Hyperparameter Optimization

Here we will optimize the standard likelihood function (log) using L-BFGS-B implemented in *sci-py*. I will spare you the gory details, and I'm not 100% sure I'm correct in the calculation of my gradient, but below is a plot of some random data with the following functional form:

$$f_0(x) = -X_1 + U(-10^{-2}, 10^{-2})$$

$$f_1(x) = -X_1 + U(-10^{-1}, 10^{-1})$$

Where we will (i) only use X_1 in $f(x)$ for simplicity and where if the experiment has a 0 in it's first column it'll be assigned to $f_0(x)$ and if it has a 1 in the first column it'll be assigned to $f_1(x)$. To help, I will bifurcate the data such that we have equally space data points in X_1 and one data set will sit on the right side of X_1's design space and the other will occupy the remaining left side.

As you can see both implementations successfully recapitulated the noisy linear trend in red, while Botorch's regular GP (only having access to the highest fidelity model $l = 0$ did not do as good. Of course what is the value in this? As I said in the top of the post, if we have multiple types of sensors attempting to measure the same thing, we can statistically fuse their data across quality l and space x. So our toy problem might represent a type of sensor that is common in one area and not in another.

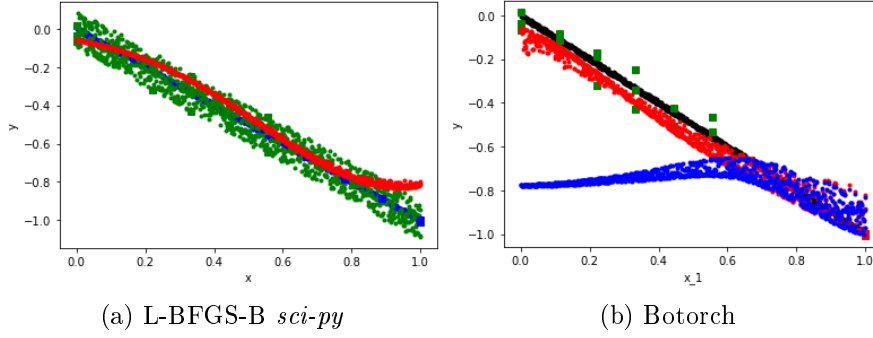(a) L-BFGS-B *sci-py* (b) Botorch

Figure 1: Multi-Fidelity GP | **left** my implementation in python (see code) for prediction of $f_0(x)$ in red **right** Botorch implementation in python for prediction of $f_0(x)$ in red, regular GP using only 0th data in blue, for both plots blue and green squares are the $l = 1$ and $l = 0$ fidelity data respectively

An important thing to note here is that we have treated our problem $f_0(X)$ as being a measure of the true underlying value (or at least the closest we can get) while $f_1(x)$ was merely an approximation. This hierarchy need not be the case as we can predict whatever fidelity we wish by changing the first column in our data l. There are other models that treat all information sources equally (rather than being deviations from an underlying process) and we may discuss them in future.

Until then we are done! This week was a real fun one but quite difficult. I'll try to do something lighter for next week but haven't decided. See ya!